

Published by
The Hong Kong College of Family Physicians
7th Floor, HKAM Jockey Club Building,
99 Wong Chuk Hang Road, Hong Kong
Tel 2528 6618 Fax 2866 0616
Website: <http://www.hkcfp.org.hk>

EDITOR

Dr. E Y Y Tse 謝婉儀醫生

DEPUTY EDITORS

Dr. D Owens 歐德維醫生
Dr. F C T Lee 李長智醫生
Dr. K K Ng 吳國強醫生

EDITORIAL BOARD MEMBERS

Dr. S T Borwein
Dr. N Y Chan 陳迺賢醫生
Dr. D V K Chao 周偉強醫生
Dr. S Chen 陳紹德醫生
Dr. L P K Cheng 鄭佩君醫生
Dr. M K Cheung 張文娟醫生
Dr. K W Ching 程錦榮醫生
Dr. R W M Chow 周偉文醫生
Dr. A A T Chuh 許晏冬醫生
Dr. X Fu 付希娟醫生
Dr. K K L Kung 龔敬樂醫生
Dr. B K M Kwan 關嘉美醫生
Dr. K H Kwok 郭冠雄醫生
Dr. C L K Lam 林露娟醫生
Dr. M Lam 林民醫生
Dr. T K Lam 林大鈞醫生
Dr. T P Lam 林大邦醫生
Dr. W W Lam 林永和醫生
Dr. Y Lam 林遠醫生
Dr. H L Lau 劉浩濂醫生
Dr. S M Lee 李詩眉醫生
Dr. J Liang 梁峻醫生
Dr. C C Y Ng 伍志然醫生
Dr. C B Ng 吳進坡醫生
Dr. T K Ng 吳梓江醫生
Dr. R W M Pau 包偉民醫生
Dr. A T Y Tang 鄧天旭醫生
Dr. K K Tse 謝國基醫生
Dr. K K L Tsim 詹觀蘭醫生
Dr. W S Tsui 徐詠詩醫生
Dr. K W Wong 黃家華醫生
Dr. W C W Wong 黃志威醫生
Dr. Y K Yiu 姚玉筠醫生

BUSINESS MANAGER

Dr. M B L Kwong 鄭碧綠醫生

ADMINISTRATIVE EXECUTIVE

Ms. T S K Lee 李小姐女士

EDITORIAL BOARD ADVISERS

Prof. H Chiu 趙鳳琴教授
Prof. J Dickinson 狄堅信教授
Prof. T Dixon 迪辰教授
Prof. W E Fabb 費伯威教授
Prof. S T Fan 范上達教授
Prof. C B Hazlett 希士偉教授
Prof. A J Hedley 賀達理教授
Prof. W W K King 金永強教授
Prof. C R Kumana 顧崇仁教授
Prof. C P Lau 劉柱柏教授
Dr. P C Y Lee 李仲賢醫生
Prof. P C Leung 梁秉中教授
Prof. K Y Mak 麥基恩教授
Prof. W C G Peh 白振源教授
Prof. G Tang 鄧惠瓊教授
Prof. C A Van Hasselt 尹懷信教授
Prof. J Woo 胡令芳教授
Prof. T W Wong 黃子惠教授
Prof. C Y Yeung 楊執庸教授
Prof. R T T Young 楊紫芝教授

STATISTICAL CONSULTANTS

Prof. J Lau 劉德輝教授
Dr. I Lauder 羅德賢博士

Printed and designed by
Printhouse Workshop, Hong Kong

Assessment – We don't like it but we cannot do without it

Cindy L K Lam 林露娟

One has to go through no less than 100 formative or summative assessments from primary school to university graduation in order to qualify as a doctor. Added to these are the Conjoint Fellowship Examination and Exit Examination in order to become a specialist in family medicine. Although we are, as described by the Chinese saying “身經百戰 (has experienced 100 wars)”, experts in assessments, we still find them stressful and some may even find them demoralizing. Assessment is one of those things in life, like taxation, that we don't like but we cannot do without.

Assessment purpose

Assessment is an indispensable part of education and training to ensure that a candidate has achieved the required standard for the award of a qualification, promotion, progress to the next stage of training, professional development or quality assurance. Summative assessment that determines a pass or fail against a pre-set standard is the commonest method used in statutory or professional qualifying examinations. It is the best available method to assure that the minimum required level of competence has been reached. Formative assessment uses repeated measurements over a period of time, which is more suitable for monitoring the progress of training and for appraisal.

Ideally professional assessments should assess the performance of the candidate, which is the highest level of competence, as described in Miller's Pyramid of Competence.¹ The foundation of the pyramid is knowledge (know what), on which one builds skills (know how) and acquires the ability (show how), and the ultimate outcome is performance (put to practice). Different methods have to be used to assess different levels of competence. The widely used multiple choice questions mainly tests knowledge, while the objective structure clinical examination (OSCE) is a popular method to test skills and ability.² The assessment of performance requires direct observation of practice

HK Pract 2007;29:177-179

Cindy L K Lam, MBBS (HK), MD (HK), FRCGP, FHKAM (Family Medicine)
Clinical Associate Professor,
Family Medicine Unit, The University of Hong Kong.

Correspondence to : Dr Cindy L K Lam, 3rd Floor, Ap Lei Chau Clinic, 161 Main Street, Ap Lei Chau, Hong Kong.

is not as commonly done as it should be because it is very challenging to organize and standardize. Recent developments in the methods of practice-based assessment have enabled performance assessment to be done validly, objectively and reliably. The Leicester Assessment Package of consultations³ used in the Exit Examination of the Hong Kong College of Family Physicians (HKCFP) is an example of such a method.

Irrespective of the method, an assessment must be valid, reliable, fair and discriminatory in order to serve its purpose.

Assessment validity and reliability

Validity means the test measures what it intends to measure. Competence, like beauty, is a latent variable that cannot be directly measured. It can only be assessed indirectly through the use of correlated indicators that are observable. These indicators become the content (questions) of an assessment. In order to serve the purpose, the indicators should be important, relevant, representative and adequate of the target competence being tested for. Reliability is a measure of consistency and reproducibility, which is a function of sample size in that the larger the sample of domains, questions, examiners, and methods, the more reliable are the results.⁴ It is important to note that reliability is an essential but insufficient requirement of a test. A very reliable test that is not valid is completely useless. For example, multiple choice questions are very reliable because they include a large sample but they have low validity as tests of skills and performance.

Assessment criteria and standards

An assessment needs to be fair in that all candidates with similar competence should achieve similar outcomes. This is easy with multiple choice questions for which model answers can be set, but it is not as straight forward for other assessment methods. Criteria that define the desirable attributes or behaviour are required for OSCE and other tests on skills and attitudes. Generic criteria that are independent of the case content are needed for practice-based assessment.

The ultimate goal of a professional qualifying assessment is to discriminate candidates who have passed the standard (required level of competence) from those who have not. The standard for a professional assessment

should be absolute and based on pre-determined criteria, such as those used in the HKCFP Conjoint Fellowship and Exit Examinations. For example, the passing threshold for the HKCFP Exit Examination is 65% that is defined as “consistently demonstrates capability in almost all components to a high standard and a satisfactory standard in all”. Some assessments, e.g. Hong Kong Certificate of Secondary Education Examination, use relative standards with a fixed passing rate and the outcome of a candidate depends on the standard of others. This is not appropriate for professional qualifying examinations because the competence level required of a professional should be absolute rather than normative.

One challenge with absolute standard setting is the effect of the level of challenge of the examination questions on the scores that candidates can achieve. Different methods have been developed to adjust for this.⁶ One of them is the Contrasting Groups method in which the examiner gives a global rating on the pass level for the candidate in addition to the score against the assessment criteria for the question.⁶ The cut-off score of each question is then determined statistically from the distribution of the scores of borderline candidates. This method is used in the OSCE of the HKCFP Conjoint Fellowship Examination to ensure fairness of the assessment.

Conclusions

Professional competence is multi-dimensional; it is not possible to have one single test that can validly and reliably measure all the required indicators. Multiple methods have to be used to determine whether a candidate has achieved the required level of competence. An adequate number of questions and examiners should be used to improve the reliability. Clearly defined assessment criteria are essential to ensure fairness. Standards should be set at an absolute level to assure the quality of the profession.

Assessment is a complex form of measurement. All measurements reflect not only the true result (the candidate's competence) but also variations that are related to the assessment content and context, examiners and other unknown factors. Research has shown that candidate factors explain no more than 50% of the variance in the results.⁷ No assessment is perfect; we can only try to make it as good as possible through a rational choice of methods and questions, fair setting of criteria and standards, and adequate examiner training. ■

References

1. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990;65(Suppl):S63-S67.
2. Newble D. Techniques for measuring clinical competence: objective structured clinical examinations. *Med Educ* 2004;38:199-203.
3. Fraser RC, McKinley RK, Mulholland H. *Assessment of consultation competence in general practice: the Leicester Assessment Package in Approaches to the assessment of clinical competence. Part 1.* In: Harden RM, Hart IR, Mulholland H (ed). Dundee: Centre for Medical Education. 1992:192-198.
4. Downing SM. Reliability: on the reproducibility of assessment data. *Med Educ* 2004;38:1006-1012.
5. Lau JKC, Fraser RC, Lam CLK. Establishing the content validity in Hong Kong of the prioritized criteria of consultation competence in the Leicester Assessment Package (LAP). *HK Pract* 2003;25:596-602.
6. Norcini JJ. Setting standards on educational tests. *Med Educ* 2003;37:464-469.
7. Van der Vleuten CPM. The assessment of professional competence: developments, research and practical implications. *Advances in Health Sciences Education* 1996;1:41-67.